



# The Cortical Underpinnings of Context-based Memory Distortion

## Citation

Aminoff, Elissa, Daniel L. Schacter, and Moshe Bar. 2008. The Cortical Underpinnings of Context-based Memory Distortion. *Journal of Cognitive Neuroscience* 20(12): 2226-2237.

## Published Version

<http://dx.doi.org/10.1162/jocn.2008.20156>

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:3199017>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# The Cortical Underpinnings of Context-based Memory Distortion

Elissa Aminoff<sup>1,2</sup>, Daniel L. Schacter<sup>1</sup>, and Moshe Bar<sup>2</sup>

## Abstract

■ Everyday contextual settings create associations that later afford generating predictions about what objects to expect in our environment. The cortical network that takes advantage of such contextual information is proposed to connect the representation of associated objects such that seeing one object (bed) will activate the visual representations of other objects sharing the same context (pillow). Given this proposal, we hypothesized that the cortical activity elicited by seeing a strong contextual object would predict the occurrence of false memories whereby one erroneously “remembers” having seen a new

object that is related to a previously presented object. To test this hypothesis, we used functional magnetic resonance imaging during encoding of contextually related objects, and later tested recognition memory. New objects that were contextually related to previously presented objects were more often falsely judged as “old” compared with new objects that were contextually unrelated to old objects. This phenomenon was reflected by activity in the cortical network mediating contextual processing, which provides a better understanding of how the brain represents and processes context. ■

## INTRODUCTION

Memory for past events is not perfect; it typically involves forgetting, adding to, or distorting details of an actual episode (Loftus, 2003; Schacter, 1999; Bartlett, 1932). One common type of memory distortion is referred to as “false recognition”: an incorrect claim to have seen or encountered a novel object or an event (Slotnick & Schacter, 2004; Roediger & McDermott, 1995; Underwood, 1965). We hypothesize that one source for such memory errors is related to the coactivation of contextually related objects in memory.

In everyday life, we do not encounter objects in isolation but rather, they are embedded in a context with other objects that frequently share the same context. For example, when walking into a bedroom, one typically encounters a bed, a dresser, a mirror, and an alarm clock nearby. Similarly, when encountering a cluster of balloons, we often expect a celebratory event that frequently involves gifts, a cake, and candles. Our experience with such typical settings creates in memory collections of contextually associated objects, termed “context frames” (Bar, 2004; Bar & Ullman, 1996). This study examines how the context-based coactivation of such associations, and the corresponding neuronal activity of the context processing regions of the brain, contributes to false recognition of common objects. In other words, whether the exposure to a strong contextual object (e.g., a traffic light) affects our memory such that we later falsely believe that

a contextually related object (e.g., parking meter) has been presented previously although it has not. Addressing this question will help to illuminate the cortical mechanisms mediating the phenomenon of false memory, and potentially shed more light on our understanding of the organization of contextual associations in the brain.

Prior studies have shown an increased likelihood of incorrectly recalling an object that is contextually related to an object previously encountered in a scene, compared with an unrelated object. For instance, Brewer and Treyens (1981) left participants in an office for 10 min, after which the participants were moved out of that office to another room, where a surprise memory recall test was administered. Participants falsely “recalled” items that were contextually congruent with an office more often than items that were not contextually congruent. Miller and Gazzaniga (1998) demonstrated similar contextual effects. Participants were shown a series of Norman Rockwell illustrations of typical scenes (e.g., a beach) and were subsequently tested on their memory for the items within the scenes. Participants were more likely to produce false alarms to items that were contextually congruent with the scenes (e.g., a beach ball) compared with items that were unrelated (e.g., a chalkboard).

Both of these previous studies relied on entire scenes to generate false memory. Therefore, in addition to our novel examination of the neural correlates of context-based false recognition, the present study goes beyond previous reports by investigating the actual mechanism by which individual objects can evoke an activation

---

<sup>1</sup>Harvard University, <sup>2</sup>Harvard Medical School

pattern in which the coactivation of contextual related objects, not present at the time of encoding, can lead to subsequent false recognition.

The neural origins of contextually related false memories are largely unknown. Recent work (Aminoff, Gronau, & Bar, 2007; Bar, 2004; Bar & Aminoff, 2003) has revealed three main cortical areas that mediate contextual processing: the parahippocampal cortex (PHC), the retrosplenial complex (RSC), and a third focus in the medial prefrontal cortex (MPFC) observed under certain task demands (Bar, 2007; Bar, Aminoff, Mason, & Fenske, 2007). Using functional magnetic resonance imaging (fMRI) across several studies, these regions were found to be selectively activated when participants viewed objects with strong contextual associations (e.g., a traffic light, strongly associated with a street context; or a baby bottle, strongly associated with the context of a baby) compared with objects with weak contextual associations (e.g., a camera, not specifically associated with a particular contextual setting). We concluded that the PHC, the RSC, and the MPFC constitute a network subserving contextual associations (a context network “localizer” is available at <http://barlab.mgh.harvard.edu/ContextLocalizer.htm>). The characterization of a neural system for contextual processing provides a basis for exploring the neural activity associated with contextually related false memories. Specifically, we hypothesize that the coactivation of these contextual associations at the time of encoding is responsible for subsequent false recognition of related contextual objects. To examine the neural origins of such false recognition, we used a subsequent memory paradigm in which neural activity at the time of encoding is related to later remembering or forgetting (Brewer, Zhao, Desmond, Glover, & Gabrieli, 1998; Wagner et al., 1998). The subsequent memory paradigm was previously applied to false recognition in a study by Gonsalves and Paller (2000), who used it to examine neural events at encoding associated with perceiving versus imagining an object as a function of true or false recognition on a later test. Using this paradigm allows us to examine encoding activity as a function of later recognition accuracy on an item-by-item basis (Gonsalves & Paller, 2000). The subsequent memory method has been successfully used in prior work to examine the origins of different types of memory errors, including source monitoring (Gonsalves et al., 2004; Gonsalves & Paller, 2000), emotional content influences on source monitoring (Kensinger & Schacter, 2005), and when general information, rather than specific information, is retained in memory (Garoff, Slotnick, & Schacter, 2005).

Our main hypothesis is that increased activity in the cortical network that subserves contextual processing will be predictive of subsequent false recognition of objects that are contextually related to the presented object. According to this hypothesis, greater activity elicited in the context cortical network would indicate an increased activation of contextually related objects,

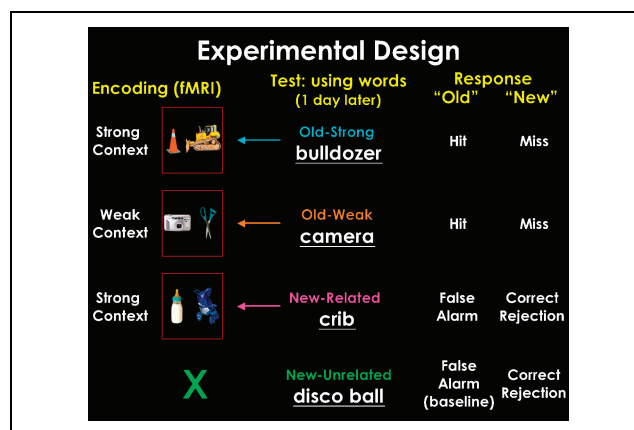
and this increased activation of contextually related objects would, in turn, lead to increased likelihood of false recognition of these related items.

In the present study, participants were scanned using fMRI during the encoding phase of the experiment. We designed a novel paradigm where two common visual objects were presented simultaneously on each encoding trial, and participants were asked to “put the two objects together into a context and to press a button based on how many objects they associated with the context” on a 4-point scale (see Methods). The pairs of objects were of one of two types: either two strong contextual objects that belonged to the same context (e.g., a bulldozer and a construction cone) or two weak contextual objects that were not related to each other (e.g., a camera and a pair of scissors; Figure 1). A day after the encoding phase, participants performed, outside of the scanner, an old–new recognition memory test in which they were presented with words depicting objects from one of four types of categories: (1) strong contextual objects that were presented at encoding (e.g., bulldozer; old–strong item); (2) weak contextual objects presented at encoding (e.g., camera; old–weak item); (3) strong contextual objects *related to the context* presented at encoding, but that did not actually appear at encoding (e.g., crib, with relation to baby bottle and stroller presented at encoding; new–related lure item); and (4) novel objects (both strong and weak contextual) not related to any object or context presented at encoding (e.g., disco ball; new–unrelated baseline).

## METHODS

### Participants

Twenty-five participants were scanned in this experiment. Nine participants were excluded from the analysis based



**Figure 1.** Experimental design. Participants were scanned while presented with pairs of objects that either had strong contextual associations or weak contextual associations. The following day, participants were tested for their memory with words. Based on their response, encoding trials were defined as trials that led to a hit, miss, false alarm, and correct rejection.

on various criteria (see Data Analysis). The remaining 16 participants consisted of 6 women (15/16 right hand dominant), with a mean age of 26.75 years ( $SD = 3.67$ ). All participants had normal or corrected-to-normal vision. Informed written consent was obtained from each of the participants prior to the scanning sessions. All procedures were approved by Massachusetts General Hospital Human Studies Protocol number 2001-001754.

## Stimuli

Visual objects used in the experiment were either strongly related to a particular context (SC), or weakly associated with many contexts (WC) (see Figure 1). Objects were rated as either SC or WC based on previous pilot surveys as described in Bar and Aminoff (2003). There were a total of 84 different contexts presented to the participant. There were three different objects associated with each context: two “key” objects, rated as most typical in the context, and one “relevant” object that was not highly associated with the context. Each participant at encoding saw one key object and one relevant object from each context on a given a trial. The purpose of this balancing was that one key object would be shown at encoding, and the other at test as a contextual lure item. Key objects were balanced between participants. There were a total of 252 SC objects used in this experiment, where only 168 of the SC objects were shown at encoding. There were also a total of 144 WC objects used at encoding. Fifty-seven new-unrelated objects were used as new items at test to obtain a baseline false alarm rate.

## Procedure

Participants were scanned while they viewed photographs of everyday objects on a gray background. On each trial, two objects were presented side by side in the center of the screen. Each individual picture was  $9^\circ$  of visual angle; the two pictures together spanned a visual angle of  $20^\circ$ . The pictures of objects were presented on a black screen. Each picture pair was presented for 1500 msec and there was a 1500-msec interstimulus interval. Picture trials were intermixed with fixation trials in a predetermined order to maximize efficiency and accuracy in extracting the hemodynamic response function (order was created by the function `optseq`, part of the FreeSurfer toolbox; <http://surfer.nmr.mgh.harvard.edu/optseq/>). There were a total of 156 picture trials (84 SC trials and 72 WC trials) and 126 fixation trials spread over three functional runs.

The task for participants was to try to create a context between the two objects, and to press a button based on how many objects they associate with the context. Participants pressed “1” if they associated many objects with the context; “2” if they associated just a few additional items with context; “3” if they could take the two objects presented and put them into a context

together but did not associate any other objects with the context; and “4” if they could not even put the two objects presented in a context together.

The next day, participants returned for the testing period of the experiment. In the test, a word was presented and the participant determined whether the word was presented the day before as a picture, and thus, is “old”; or if the word was not a picture presented the day before and therefore is “new.” If the participants decided the word was “old,” they were asked to make a further judgment of whether they vividly remember seeing the picture (i.e., “remember”), or if they just had a feeling of knowing that the word was presented as a picture the day before (i.e., “know”). Each word corresponded to one of four conditions: a strong contextual item presented the day before as a picture (strong-old); a weak contextual item presented the day before (weak-old); a strong contextually item *related* to a contextually related pair at encoding (strong-lure); and a new item unrelated to any of the pictures presented at encoding (new-unrelated). The contexts were split such that half the contexts ( $n = 42$ ) were presented with a strong context old item, and the other half ( $n = 42$ ) were presented with a strong context lure item. Contexts were balanced between participants. Half the weak context items presented at encoding were presented at test as a weak-old item ( $n = 36$ ). There were 57 new-unrelated trials.

## Imaging Parameters

The participants engaged in the encoding phase while whole-brain fMRI scans were collected on a 3-Tesla Siemens Allegra head-only scanner using a gradient echo-planar imaging sequence ( $TR = 3000$  msec,  $TE = 25$  msec, flip angle =  $90^\circ$ ). The acquired slices were axial, parallel to the anterior commissure–posterior commissure line (33 slices, 3 mm, 1 mm skip). Each participant participated in a series of anatomical scans as well as three functional scans.

## Data Analysis

The data from seven participants that were originally scanned were excluded from the analysis because we used a criterion of at least 10 observations per participant in the strong false alarm condition to provide sufficient numbers of trials for the fMRI analysis. Two participants were excluded from the analyses because they did not show activation of the context network indicated by comparing strong to weak trials. The remaining 16 participants were averaged in the group analysis.

Functional data were analyzed using the FreeSurfer analysis tools. Data from individual fMRI runs were first corrected for motion using the AFNI package (Cox, 1996) and spatially smoothed with a Gaussian full-width, half-maximum (FWHM) filter of 5 mm. The intensities for all runs were then normalized to correct for signal intensity

changes and temporal drift, with global rescaling for each run to a mean intensity of 1000.

Performance on the memory test the day after the scanning session was used to back-sort the trials at encoding. Thus, each trial at encoding was defined by the performance at test. Signal intensity for each condition was then computed and averaged across runs. A finite impulse response model was used for the analysis. To account for intrinsic serial correlation in the fMRI data within participants, we used a global autocorrelation function that computes a whitening filter (Burock & Dale, 2000). The data were then tested for statistical significance and activation maps were constructed for comparisons of the different conditions. Both group-average activation maps as well as regions of interest (ROIs) are random effect analyses.

### *Cortical Surface-based Analysis*

Once the data from all trials were averaged, the mean and variance volumes were resampled onto the cortical surface for each participant. Each hemisphere was then morphed into a sphere in the following manner (Segonne et al., 2004; Fischl, Liu, & Dale, 2001; Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999). First, each cortical hemisphere was morphed into a metrically optimal spherical surface. The pattern of cortical folds was then represented as a function on a unit sphere. Next, each individual participant's spherical representation was aligned with an averaged folding pattern constructed from a larger number of individuals aligned previously. This alignment was accomplished by maximizing the correlation between the individual and the group, while prohibiting changes in the surface topology and simultaneously penalizing excessive metric distortion (Fischl, Sereno, Tootell, & Dale, 1999).

### *Region-of-interest Analysis*

The ROIs for this study were chosen by the results of the strong versus weak contrast. The PHC ROIs as well as the MPFC RH were included as well due to a priori hypothesis of the involvement in contextual processing (Bar, 2004). The PHC, RSC, and MPFC ROIs were defined structurally. The structural constraint of the PHC (encompassing the collateral sulcus and the parahippocampal gyrus) was based on a hand labeling of different brain structures for each participant. The PHC was defined using procedures elaborated in Insausti et al. (1998) and Reber, Wong, and Buxton (2002). The RSC was hand-labeled on each individual using a structural constraint based on anatomical landmarks of the calcarine sulcus, the parieto-occipital sulcus, the corpus callosum, and the posterior cingulate sulcus (refer to Figure 2). The MPFC was defined as anterior to the corpus callosum, and in front of and below the cingulate sulcus. The lateral parietal cortex (LP) and the occipito-temporal sulcus (OTS)

ROIs were defined functionally. The anatomical location of the LP and OTS ROIs were defined by the cluster of activity significant in the strong versus weak contrast from the group analysis and then projected back to each individual's brain. For all ROIs, a functional constraint was used either by selecting the subset of voxels within each of these labels which demonstrated a significant effect of context (i.e., significant in the strong versus weak contrast) when examining the activation related to false recognition, or by any component of the task, as revealed by the main effect (i.e., all vs. fixation contrast) when examining the activation related to hits and misses. All of the voxels that met these constraints were then averaged, allowing the contrasts of interest to be computed across the resulting time courses. An outlier analysis was performed for each ROI, on each individual participant, such that the signal was averaged across conditions. If a participant had an average signal that was above the statistical threshold of 2.5 standard deviations from the mean of the group, they were considered an outlier and removed from the analysis for that specific ROI. In only the right RSC was an outlier found and is noted within the results. A one-way repeated-measures analysis of variance (ANOVA) was performed for experimental conditions on the mean percentage of peak signal change calculated for each condition.

## **RESULTS**

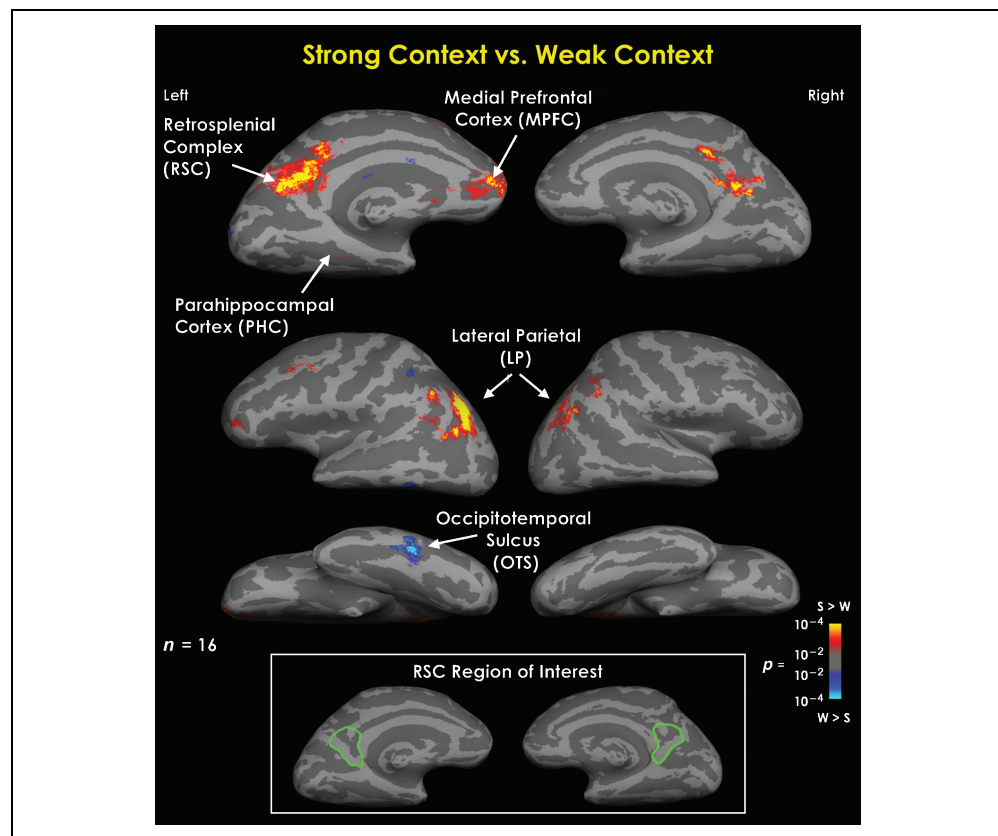
To test our hypothesis, we first examined whether context indeed affects old–new recognition performance, which would be indicated by a significantly higher false alarm rate for the new–related lures compared with the new–unrelated baseline items. Second, to examine the neural regions that predict the occurrence of subsequent contextually related false recognition, we compared neural activity at encoding between the strong contextual pairs that led to a subsequent false alarm and the strong contextual pairs that led to a subsequent correct rejection of the new–related lure items. We hypothesized that regions that mediate the processing of contextual associations would demonstrate increased fMRI signal for those encoding trials that lead to subsequent false alarms compared with trials that lead to subsequent correct rejections. Such a result would support an account of false recognition whereby such errors stem from context-related coactivations.

### **Behavioral Results**

#### *Encoding Task*

As noted earlier, the encoding task required participants to respond on a 4-point scale, where “1” reflected many associated objects with the context of the pair and “4” reflected an inability to relate the two objects presented to a shared context. The average response in the strong contextual trials was 1.43 ( $SD = .37$ ), whereas the average

**Figure 2.** Statistical parametric maps representing the difference between viewing a pair of objects that strongly relates to a specific context and viewing a pair of objects that does not associate with any context, or weakly associates with many. This is a random effects average of 16 participants. This contrast resulted in differential activation in the retrosplenial complex (RSC), the medial prefrontal cortex (MPFC), regions of the lateral parietal cortex (LP), and the occipito-temporal sulcus (OTS). The results of this analysis were used to define the ROIs in order to examine the relation between contextual processing in the brain and false recognition of contextually related objects. Bottom insert box delineates the region within the medial parietal cortex included in the RSC region of interest (outlined in green).



response for the weak contextual trials was 3.27 ( $SD = .51$ ). This difference in relating the two objects to a larger context was statistically significant [ $t(15) = 15.4, p < .001$ ], validating our initial distinction between strong and weak contextual objects pairs. There was also a significant reaction time difference in the encoding task when comparing the strong contextual trials (1.42 sec) to the weak contextual trials [1.7 sec,  $t(15) = 4.26, p < .001$ ]. We attribute this reaction time difference to the difficulty of determining whether the weak contextual pair of objects fit into a context together and the ease to which a context was found for the strong contextual trials.

### Recognition Task

A significant effect of contextual associations was obtained for the likelihood of false recognition to a related item, relative to unrelated new item (see Table 1). Specifically, participants made 18% more false recognition responses (i.e., rating a new item as “old”) for contextually related lure (new-related) items compared to novel, unrelated items [new-unrelated;  $t(15) = 6.75, p < .001$ ]. A similar contextual effect was also obtained for true recognition of old items in the strong (old-strong) versus the weak (old-weak) contextual condition. Namely, participants recognized 21% more strong than weak old contextual items [ $t(15) = 6.95, p < .001$ ].

Although no significant differences in reaction times were found at test, participants were faster to respond

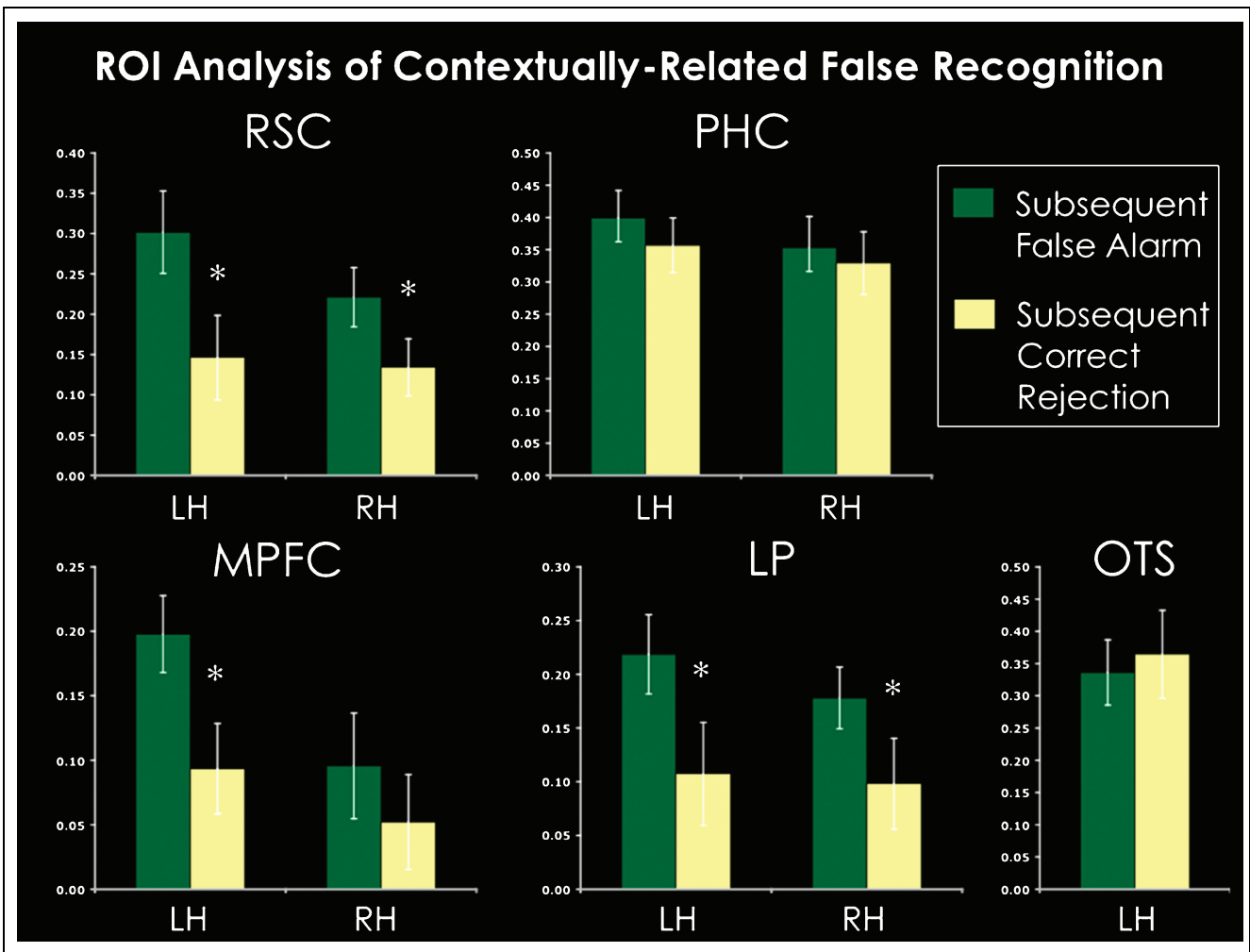
in the encoding phase to the strong contextual trials that later led to a false alarm (1.38 sec) compared to those trials that led to correct rejections (1.43 sec), hits (1.43 sec), and misses (1.53 sec). This difference only reached marginal significance at a two-tailed comparison with correct rejections [ $t(15) = 2.00, p < .06$ ] and hits [ $t(15) = 1.84, p < .09$ ]. There was no significant reaction time differences between the weak contextual trials that resulted in a subsequent hit (1.69 sec) versus a miss [1.71,  $t(15) < .7$ ].

An ANOVA demonstrated a marginal effect of response choice (i.e., the 4-point scale determining the

**Table 1.** Recognition Memory Performance

Condition	Recognition Memory	
	Response	
	“Old”	“New”
Old-Strong	65 ( $R = 55$ )	35
Old-Weak	44 ( $R = 43$ )	56
New-Related ( <i>lure</i> )	45 ( $R = 33$ )	55
New-Unrelated ( <i>baseline</i> )	27 ( $R = 31$ )	73

Average of 16 participants. Numbers represent the percentage of trials of each condition that were perceived as old and new at test. Within the old trials, the percentage of “remember” responses compared to “know” responses are in parentheses.



**Figure 3.** ROI analysis for regions that process contextual associations. Conditions shown are the encoding data for the strong contextual trials that, at test, participants had to determine if a related lure item was old (i.e., a false alarm) or new (i.e., a correct rejection). Error bars represent one standard error. \* $p < .05$ .

amount of objects they associated with the pair) at encoding on subsequent memory performance of the strong contextual trials [ $F(45, 3) = 2.58, p < .07$ ]. There was no significant difference in the rating response at encoding between the weak hits (3.24) compared to the weak misses [3.3,  $t(15) < 0.53$ ].

## fMRI Results

### Localizing the Context Network

The initial stage of the fMRI analysis was to localize areas within the cortex that show differential activity specific to contextual processing. To accomplish this objective, trials in which the participants associated the pair of objects with a context with other objects (i.e., a response of “1” or “2”) were compared with the trials in which the participant did not associate the pair of objects with any context and therefore were considered a weak context, or a noncontextual trial (i.e., a response

**Table 2.** Statistical  $t$  and  $p$  Values for ROI Analysis Comparing Encoding Data for Trials that Led to Subsequent False Alarm with Data for Trials that Led to Subsequent Correct Rejections

ROI	$n$	$t$ Statistic	$p$
RSC LH	16	3.20	.005*
RSC RH	15	2.44	.028*
MPFC LH	16	2.88	.011*
MPFC RH	16	.94	.36
LP LH	16	2.64	.018*
LP RH	16	2.13	.05*
PHC LH	16	.85	.4
PHC RH	16	.40	.69
OTS LH	16	-.53	.6

All  $p$  values are for two-tailed significance.

\* $p < .05$ .

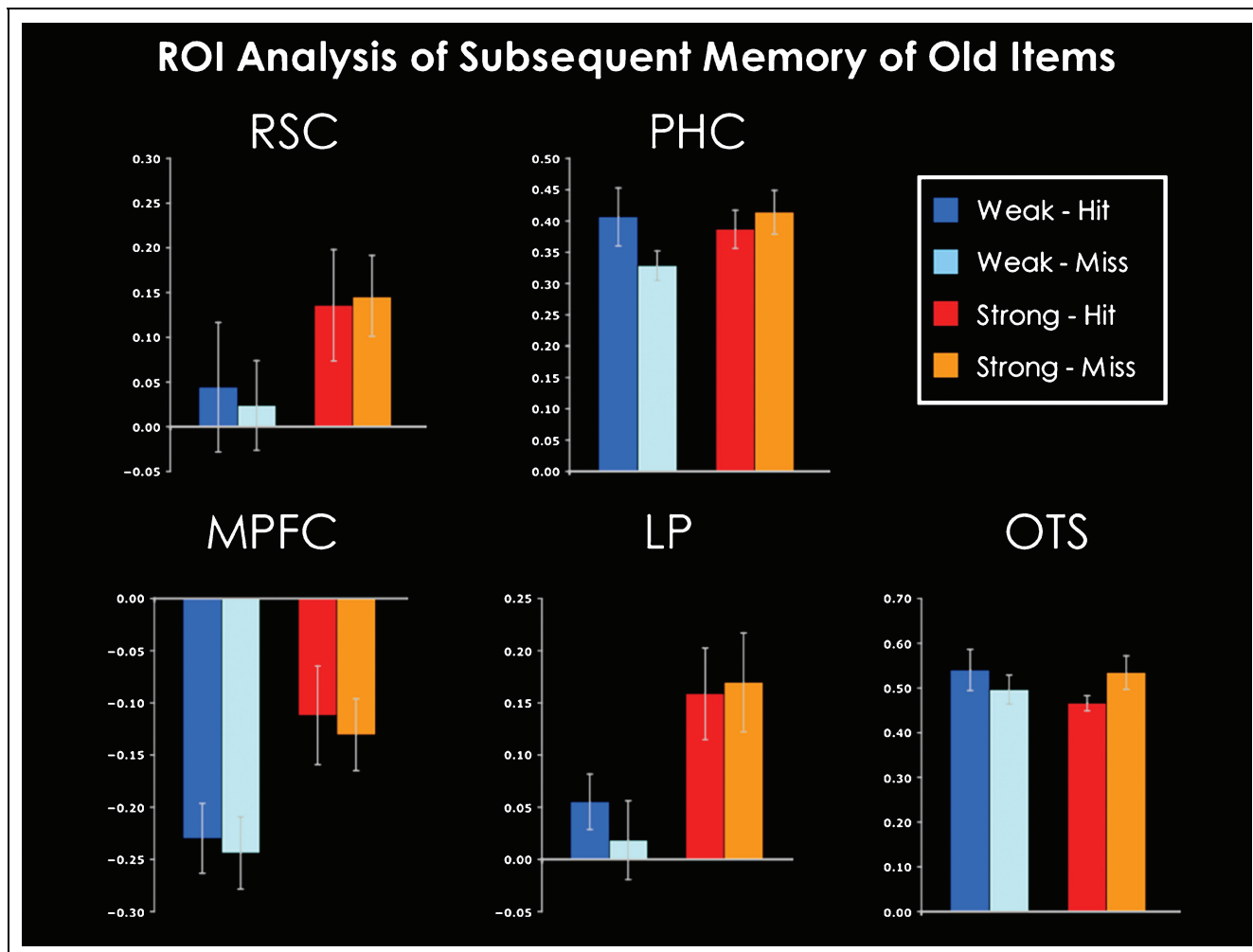


of “4”). The results of this comparison can be seen in Figure 2. The differential activity exhibited from this contrast revealed four main sites of activation: the bilateral RSC [which includes the retrosplenial cortex, and parts of the precuneus and posterior cingulate; Talairach: LH (−6, −56, 18), RH (4, −47, 15)], the left MPFC [Talairach: LH (−9, 41, −6)], the bilateral lateral parietal [LP, including areas of the inferior parietal lobule, and the supramarginal gyrus; Talairach: LH (−44, −66, 36) and (−45, −42, 56), RH (26, −64, 52) and (46, −51, 42)], and an area within the left occipito-temporal sulcus [OTS; Talairach: (−56, −51, −10)]. Out of these four sites, all were more active for context trials compared to no-context trials except for the OTS, which showed greater activity for the no-context trials compared to context trials. This analysis was used to determine the ROIs in order to examine what activity at encoding is related to contextual false recognition. Based on our previous re-

search on cortical areas mediating contextual associations, we included a PHC ROI and a left hemisphere MPFC, both of which have been shown to be involved in contextual processing. We discuss later the implications of why PHC activity was not evident in the contrast comparing context to no-context pairs.

#### *False Alarm vs. Correct Rejection of New-Related Items*

ROI analyses were run on all the regions that exhibited differential activity in context compared to the no-context trials as well as the PHC, to investigate whether activation at encoding in these regions related to subsequent false recognition of contextually related items. Results are shown in Figure 3 and Table 2. Descriptions of how the ROIs were labeled in each individual can be found in the Methods section. In each ROI, we compared activity at encoding elicited for those trials that led to a false



**Figure 4.** ROI analysis for regions that process contextual associations. Trends were similar in both hemispheres, and thus, collapsed in the figure for simplicity. Conditions shown are the encoding data for the subsequent memory of an old item. Items were identified as strong or weak based on the initial presentation of the object: either in a strong contextual pair (strong) or in a pair of weak contextual objects (weak). Items correctly identified as old were considered a hit or if the old item was misidentified as new, it was considered a miss. Error bars represent one standard error.



**Table 3.** ROI Statistical Values of Subsequent Memory of Old Items

ROI	Context	Memory	Con $\times$ Mem
RSC LH	$F(1, 10) = 3.07, ns$	$F(1, 10) = 0.3, ns$	$F(1, 10) = 3.2, p < .1$
RSC RH	$F(1, 10) = 6.37, p < .03$	$F(1, 10) = 0.06, ns$	$F(1, 10) = 0.28, ns$
MPFC LH	$F(1, 10) = 3.01, ns$	$F(1, 10) = 3.20, ns$	$F(1, 10) = 0.01, ns$
MPFC RH	$F(1, 10) = 26.58, p < .001$	$F(1, 10) = 0.48, ns$	$F(1, 10) = 0.32, ns$
LP LH	$F(1, 10) = 14.01, p < .004$	$F(1, 10) = 0.07, ns$	$F(1, 10) = 0.42, ns$
LP RH	$F(1, 10) = 7.53, p < .02$	$F(1, 10) = 1.05, ns$	$F(1, 10) = 1.93, ns$
PHC LH	$F(1, 10) = 0.91, ns$	$F(1, 10) = 0.07, ns$	$F(1, 10) = 2.11, ns$
PHC RH	$F(1, 10) = 0.45, ns$	$F(1, 10) = 5.94, p < .04$	$F(1, 10) = 3.31, p < .1$
OTS LH	$F(1, 10) = 0.26, ns$	$F(1, 10) = 0.09, ns$	$F(1, 10) = 4.81, p < .05$

Statistical  $F$  and  $p$  values for ROI ANOVA looking at the main effect of context compared to the main effect of memory (i.e., recognition of old items) for encoding data of old items.

All  $p$  values are for two-tailed significance.

recognition of a contextually related item (i.e., a false alarm) with the activity elicited for those trials that led to a subsequent correct rejection, where participants correctly identified the new item as new. Only voxels within these ROIs that demonstrated a significant differential activity for context versus no-context were included in the analysis defined on a participant-by-participant basis. In the right RSC, one participant was removed due to particularly noisy data in this specific region, such that the activity from this region averaged across all conditions was above two standard deviations than the average of the group. The bilateral RSC, the left MPFC, and the bilateral LP all demonstrated a significant effect related to false recognition (Table 2 for statistical values). In each of these regions, during encoding, significantly more activity was elicited for those trials that led to subsequent false recognition compared to those trials that led to subsequent correct rejection of a contextually related item. In the one ROI that demonstrated greater activity elicited for no-context trials than context trials, there was a nonsignificant trend for the trials that led to correct rejections to activate more than those trials that led to false recognition. The results from this analysis indicate that activity in the regions that process contextual associations, specifically the RSC, the MPFC, and the LP, are predictive of subsequent false recognition for contextually related items.

#### Hits vs. Misses

In addition to examining the role of context in mediating false recognition, we conducted a post hoc analysis for other subsequent memory effects within the cortical contextual network. The data from five participants were removed from this specific analysis of hits versus misses because of an insufficient number of miss trials using a

criterion of at least 10 observations. Removing these participants did not change the trends in the false recognition data. These ROI analyses were run on all voxels that showed significant differential activity from baseline for any task-related activity performed on a participant-by-participant basis. This method allowed us to look at subsequent memory effects for both strong contextual items and weak contextual items, such that the voxels chosen were not biased toward strong context. Results are shown in Figure 4.

An ANOVA was run on each ROI examining the main effect of context and subsequent recognition of old items in each region. The statistical values can be found in Table 3. This analysis revealed a significant main effect for context, but not for recognition of old items (i.e., hits and misses) in the RSC RH, MPFC RH, and bilateral LP. Although not significant, in the left hemisphere RSC and MPFC, the main effect of context trend was found as well. A main effect for recognition of old items was found only in the PHC RH. In each ROI, except for the OTS, the weak-miss condition elicited the least amount of activity, whereas both strong memory conditions, sometimes more for the strong-miss condition, elicited the most amount of activity. We propose that this pattern of activation is a demonstration of a contextual processing gradient, such that the most contextual processing occurs for the strong-miss and strong-hit condition, and the least, or none at all, occurs for the weak miss condition. This proposal will be elaborated in the Discussion.

## DISCUSSION

Consistent with our primary hypothesis, activity in cortical areas related to contextual processing was found to be predictive of subsequent false recognition of new

items that were contextually related to items presented at encoding. Combining behavioral and fMRI findings, we propose that this false recognition is a result of the coactivation of contextually associated information at the time of encoding.

Previously we have defined the cortical network that processes contextual associations to include the RSC, PHC, and MPFC (Bar & Aminoff, 2003). The cortical regions with differential activity related to processing contextual associations in this experiment was defined as the collection of areas showing greater activity when participants viewed pairs of objects with strong contextual associations compared with pairs of objects with weak contextual associations. The regions with differential activity include both overlap with the previously defined cortical network processing contextual associations, the RSC and MPFC, and an additional region, the LP. The PHC, although typically associated with contextual processing, did not show differential activity in this particular experiment, which we propose was a result of the unique task used here. Each region that did show differential contextual activity also showed greater activation at encoding for pairs of objects in which a related item was later falsely recognized as old, compared with when the related item was correctly identified as new. It is important to consider the possible contribution of each of these regions to the generation of contextual activation and false memories, and we elaborate on it next.

A visual context contains associative information about identities of objects that tend to share the same context, as well as the typical spatial relations between them when applicable. These contextual associations are bound together in a stored memory representation referred to as a “context frame” (Bar, 2004; Bar & Ullman, 1996). The activation of a context frame presumably results in the activation of the associations inherent to that context, which we propose is the mechanism by which subsequent false recognition of contextually related items occurs. Data from previous studies provide compelling evidence that context frames might be stored and processed in the RSC (Epstein, Parker, & Feiler, 2007; Park, Intraub, Yi, Widders, & Chun, 2007; Fenske, Aminoff, Gronau, & Bar, 2006; Bar & Aminoff, 2003).

A context frame contains prototypical information about a specific context and, accordingly, is extracted from specific exemplars. For example, regardless of whether a kitchen is stainless steel modern, or country style rustic, all typical kitchens are expected to activate the same prototypical context frame of a “kitchen.” In support of the notion that the RSC mediates context frames, we have previously shown that the RSC is not sensitive to the specific visual properties of contextual representations, and processes context in a more “gist”-like manner (Bar & Aminoff, 2003). For example, the RSC responds equally to a strong contextual object presented in isolation or within a background. The strong contextual objects activated the RSC more than the weak

contextual objects because, in both cases, the same context frame was activated, regardless of the specific visual properties of the stimulus presented. In further support of the idea that the RSC processes abstracted prototypical representations of context, and does not emphasize exact physical details, we have found the RSC to be equally active for objects strongly related to a context of a specific place (e.g., oven) and objects that are not related to a specific place (e.g., baby bottle) (Bar & Aminoff, 2003). In addition, within the realm of scene processing, evidence suggests that the RSC processes scenes on a general, or prototypical level. For example, Park et al. (2007) demonstrated that activity in the RSC was related to adding information to a scene that likely appears just beyond the borders (i.e., boundary extension); and Epstein et al. (2007) reported results suggesting that the RSC processes scenes within the context of a broader environment (e.g., a school building in relation to the campus at large) rather than what was available in the immediate sensory environment (e.g., information limited to what was presented in the picture of the building). We therefore propose that the RSC processes a general, or prototypical, representation of a context, reminiscent of our definition of context frames. Hence, the activity elicited in the RSC is a manifestation of the activation of context frames and their inherent associations. Accordingly, this activation of the associations within a context frame is the source of subsequent false recognition of contextually related items.

Contextual associations are naturally beneficial to cognition (Davenport & Potter, 2004; Bar & Ullman, 1996; Biederman, Mezzanotte, & Rabinowitz, 1982; Palmer, 1975): The coactivation of contextual associations can facilitate the recognition of other objects in the environment by providing predictions about what is likely to occur in the specific context. It has been proposed that the role of the MPFC, in particular, is to generate predictions of what to expect in the immediate environment based on analogies linking the input with memory (Bar, 2004, 2007). We propose that this occurs automatically and is the source of the MPFC activation during contextual processing at encoding. When participants were asked to put the two objects into a context and think of other objects associated with the context, the MPFC presumably was recruited to generate top-down predictions about the other objects that may appear in the same context, in cooperation with other components of the context network. The collective activation associated with these predictions has led the participants to falsely recognize contextually related items as “old.”

Contextual associations not only generate predictions but can also direct attentional resources to items in our environment (Neider & Zelinsky, 2006; Chun & Nakayama, 2000). The third area found in the context cortical network was in the LP, including parts of the inferior parietal lobule and the supramarginal gyrus. The LP has been implicated in orienting attention (Corbetta & Shulman,

2002) and in episodic memory (Wagner, Shannon, Kahn, & Buckner, 2005). Recent work demonstrated that this may be the area where long-term memory and attention interact (Summerfield, Lepsien, Gitelman, Mesulam, & Nobre, 2006). We therefore hypothesize that contextually specific processing found in this area is related to the top-down use of contextual associations to orient attention. Support for this proposal comes from a study that used a repetition priming paradigm and found similar activity in the LP related to the modulation of attentional deployment by the integration of semantic and spatial contextual information (Gronau, Neta, & Bar, 2008). This LP showed here increased activation for those encoding trials that led to false recognition. It is possible that the specific task employed here, which required participants to think of as many other objects as possible that might appear in that context, promoted orienting responses while participants performed the mental “search” of their memory. This idea might further explain why activation in this area was less pronounced in other context studies, where the task did not require such active search of associates. Future experiments will be required to test this hypothesis and to characterize the involvement of this attention-related region in contextual processing in finer detail.

Previous studies indicate that the PHC plays a central role in processing contextual associations (Aminoff et al., 2007; Bar et al., 2007; Fenske et al., 2006; Bar, 2004; Bar & Aminoff, 2003). In the present study, however, no significant PHC activity specific for strong contextual trials was found in the group average. This lack of differential activity between strong and weak trials might provide additional clues regarding the role of the PHC in contextual processing. We propose that the PHC interacts with the RSC to activate the most appropriate context frame(s) based on the physical appearance of a context. While the PHC is sensitive to the specific physical properties of the input, the RSC contains a more gist-like, prototypical representation of contexts. This role of activating contextual associations that are relevant to the current episode implies that the PHC is sensitive to the specific aspects of the immediate environment. In contrast to the RSC, which processes prototypical representations of context abstracted from the details of immediate environment, we propose that the PHC processes visually specific contextual associations that more directly relate to the immediate environment. This proposal is supported by previous work that demonstrates the PHC was sensitive to the physical properties of the stimulus, for example, whether a strong contextual object was presented by itself or within a scene (Bar & Aminoff, 2003). We also found the representations within the PHC are organized along a spatial hierarchy where more visually specific spatial representations are stored in the posterior and nonspatial representations are stored in the anterior (Aminoff et al., 2007). Epstein et al. (2003, 2007) have also provided

support for visually specific representations in the PHC by demonstrating activity related to viewpoint specific scene processing, and, furthermore, activity related to scene recognition in the PHC was limited to the immediate environment rather than a broader context. We therefore propose that this “on-line” visually specific contextual processing in the PHC reflected an attempt to retrieve contextual associations and therefore was equally active for both the strong context trials and the weak context trials. In other words, in this task, the PHC performed a similar operation of contextual activation for both strong and weak context trials, therefore not showing a difference between the conditions in its response.

Although the main focus of this study concerns the role of the context network in mediating subsequent false recognition, it is interesting to consider the subsequent recognition differences in remembering old items (i.e., hits and misses). The PHC was the ROI that yielded a significant effect of recognition memory for old items. This is in accordance with previously reports that indicate a role of the PHC in the true memories as compared to false memories (Cabeza et al., 2001). In each of the ROIs, the strong-miss condition demonstrated a trend of activating these regions the most, whereas the weak-miss condition demonstrated a trend of activating these regions the least. We propose that this pattern of results demonstrates a gradient of contextual associative processing such that the strong miss activates the related associations overly broadly such that the item-specific memory is lost; whereas the weak misses do not activate many associations, if any, and therefore have no cue to promote subsequent remembering. Strong hits activated the context regions more than weak hits due to the inherent contextual associations activated by the strong items. It was only in the PHC that the weak-hit condition activated as much as the strong context trials. We propose that this effect is related to the role of the PHC in carrying out “on-line” processing of associations. In the attempt of the PHC to activate contextual associations for the weak pair of objects, some of these items may have activated an association, although not to an extent sufficient for activating a context frame. It is possible that these associations facilitated the encoding of the item such that the item was correctly identified as old at test (i.e., weak hit). Further experiments will be needed to test the interaction of context memory and item-specific memory, specifically testing whether there are conditions where context memory overrides the item-specific memory.

It is important to characterize in more detail the exact neural mechanism that gives rise to false memory. In particular, two possibilities would have to be distinguished by future research. First, such false memory could be a result of a spreading coactivation of contextually related objects, whereby seeing a highly contextual object activates associated objects that belong in the same context frame such that later one cannot reliably distinguish

between objects that were actually presented and strongly related objects that were not presented. In a second possible mechanism, a highly contextual object activates a context gist that afterward is retrieved and helps to infer old–new responses based on a global-matching process. Previous research has shown that spreading activation and gist information each contributes to various kinds of memory distortions (for review and discussion, see Schacter & Addis, 2007; Gallo, 2006; Brainerd & Reyna, 2005; Schacter & Slotnick, 2004), but their respective roles in context-based false recognition is unknown. That the RSC has been shown to represent prototypical and non-item-specific context information supports a gist-based mechanism, whereas the fact that a gist can “bootstrap” the activation of more specific context information supports a spreading activation process.

Taken together, the results of this study provide insights into both the nature of memory distortion and the components of the cortical network that mediate contextual processing.

## Acknowledgments

We thank N. Gronau for comments on the manuscript, and D. Addis, A. Gutchess, M. Fenske, and J. Mitchell for helpful discussions. This work was supported by the James S. McDonnell Foundation #21002039, Dart Foundation, NINDS-R01 NS044319 and NS050615, NCRP-P41RR14075, NRSA-T32MH070328, and NIMH MH060941.

Reprint requests should be sent to Moshe Bar, Martinos Center for Biomedical Imaging at MGH, Harvard Medical School, 149 Thirteenth St., Rm 2301, Charlestown, MA 02129, or via e-mail: bar@nmr.mgh.harvard.edu.

## REFERENCES

- Aminoff, E., Gronau, N., & Bar, M. (2007). The parahippocampal cortex mediates spatial and nonspatial associations. *Cerebral Cortex*, 17, 1493–1503.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5, 617–629.
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11, 280–289.
- Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron*, 38, 347–358.
- Bar, M., Aminoff, E., Mason, M., & Fenske, M. (2007). The units of thought. *Hippocampus*, 17, 420–428.
- Bar, M., & Ullman, S. (1996). Spatial context in recognition. *Perception*, 25, 343–352.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14, 143–177.
- Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. New York: Oxford University Press.
- Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: Brain activity that predicts how well visual experience will be remembered. *Science*, 281, 1185–1187.
- Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, 13, 207–230.
- Burock, M. A., & Dale, A. M. (2000). Estimation and detection of event-related fMRI signals with temporally correlated noise: A statistically efficient and unbiased approach. *Human Brain Mapping*, 11, 249–260.
- Cabeza, R., Rao, S., Wagner, A., Mayer, A., & Schacter, D. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences, U.S.A.*, 98, 4805–4810.
- Chun, M., & Nakayama, K. (2000). On the functional role of implicit memory for the adaptive deployment of attention across scenes. *Visual Cognition*, 7, 65–81.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3, 201–215.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29, 162–173.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, 9, 179–194.
- Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, 15, 559–564.
- Epstein, R., Graham, K., & Downing, P. (2003). Viewpoint-specific scene representations in the human parahippocampal cortex. *Neuron*, 37, 865–876.
- Epstein, R. A., Parker, W. E., & Feiler, A. M. (2007). Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *Journal of Neuroscience*, 27, 6141–6149.
- Fenske, M. J., Aminoff, E., Gronau, N., & Bar, M. (2006). Top-down facilitation of visual object recognition: Object-based and context-based contributions. *Progress in Brain Research*, 155, 3–21.
- Fischl, B., Liu, A., & Dale, A. M. (2001). Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Transactions on Medical Imaging*, 20, 70–80.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9, 195–207.
- Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8, 272–284.
- Gallo, D. A. (2006). *Associative illusions of memory*. New York: Taylor & Francis.
- Garoff, R. J., Slotnick, S. D., & Schacter, D. L. (2005). The neural origins of specific and general memory: The role of the fusiform cortex. *Neuropsychologia*, 43, 847–859.
- Gonsalves, B., & Paller, K. A. (2000). Neural events that underlie remembering something that never happened. *Nature Neuroscience*, 3, 1316–1321.
- Gonsalves, B., Reber, P. J., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., & Paller, K. A. (2004). Neural evidence that vivid imagining can lead to false remembering. *Psychological Science*, 15, 655–660.
- Gronau, N., Neta, M., & Bar, M. (2008). Integrated contextual representation for objects' identities and their locations. *Journal of Cognitive Neuroscience*, 20, 371–388.

- Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., et al. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *American Journal of Neuroradiology*, *19*, 659–671.
- Kensinger, E. A., & Schacter, D. L. (2005). Emotional content and reality-monitoring ability: fMRI evidence for the influences of encoding processes. *Neuropsychologia*, *43*, 1429–1443.
- Loftus, E. F. (2003). Make-believe memories. *American Psychologist*, *58*, 867–873.
- Miller, M. B., & Gazzaniga, M. S. (1998). Creating false memories for visual scenes. *Neuropsychologia*, *36*, 513–520.
- Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research*, *46*, 614–621.
- Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*, 519–526.
- Park, S., Intraub, H., Yi, D., Widders, D., & Chun, M. (2007). Beyond the edges of a view: Boundary extension in human scene-selective visual cortex. *Neuron*, *54*, 335–342.
- Reber, P. J., Wong, E. C., & Buxton, R. B. (2002). Encoding activity in the medial temporal lobe examined with anatomically constrained fMRI analysis. *Hippocampus*, *12*, 363–376.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814.
- Schacter, D. L. (1999). The seven sins of memory. Insights from psychology and cognitive neuroscience. *American Psychologist*, *54*, 182–203.
- Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *362*, 773–786.
- Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron*, *44*, 149–160.
- Segonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., et al. (2004). A hybrid approach to the skull stripping problem in MRI. *Neuroimage*, *22*, 1060–1075.
- Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, *7*, 664–672.
- Summerfield, J. J., Lepsien, J., Gitelman, D. R., Mesulam, M. M., & Nobre, A. C. (2006). Orienting attention based on long-term memory experience. *Neuron*, *49*, 905–916.
- Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, *70*, 122–129.
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., et al. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, *281*, 1188–1191.
- Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, *9*, 445–453.